Wage Gaps in the American Workplace for Multiracial & Multiethnic Individuals

Jaylon Grimes

Department of Economics, The Ohio State University

Introduction

Wage gaps are a complicated problem that exist within the economic structure of the United States and many other contemporary nations. Some occur naturally as a result of differences in education, area of occupation, and position within a company or corporate structure, whereas others are the result of forms of bias and discrimination on the employer's part. While a multitude of economic research papers and experimental procedures have closely examined wage gaps between the different races/ethnicities in the United States and their potential sources, little to no work exists on how these aforementioned wage gaps impact individuals of more than one race or ethnicity, aside from a passing reference or a few sentences. As a result, this paper aims to determine whether any form of multiracial/multiethnic wage gap exists in the American workplace for any of the possible combinations (which will be described in detail in the paper's Methodology section). The secondary objective for this paper pertains to which of their two respective identities each ethnic combination's wages tend to gravitate toward, with a focus on biracial individuals of Caucasian descent.

Significance of Research

With each passing generation, the ethnic makeup of our country becomes more and more mixed, and fewer people identify as a singular race or ethnicity. Correspondingly, it is imperative that our society collectively work toward ensuring that everyone has equal academic and career opportunities regardless of their status or background. Examining multiracial groups closely and determining whether wage gaps exist for specific populations could potentially kickstart this process. Doing so could also serve as a basis for examining whether these phenomena are the result of racial discrimination or the product of some form of educational or geographic disparity. In the event that discrimination is determined to be the cause of wage disparities between different ethnic groups, or any groups for that matter, it would not only negatively impact its victims but also other employees who have no culpability in the unfair practices taking place. Companies who are perceived as discriminating against certain groups can lose millions of dollars to both potential customers losing interest and lawsuits, which can create a chain reaction ending in bankruptcy. Consequently, thousands of people can lose their jobs from the presence of wage gaps, whether they are of a discriminatory nature or not (Benjamin 2016).

Background Literature

While empirical literature on wage gaps and wage-based discrimination in the United States dates back to the 1960s, it is very difficult to determine whether survey participants identifying as multiple races or ethnicities impacted their earned income. However, they still serve as excellent references for this paper's framework and methodology. The majority of background literature for this paper consists of case studies on wage gaps around the world. Looking at wage disparities is essential to understanding the present state of the economy, as they are strong indicators of the evolution of labor force characteristics over time (Chang & Reed 2003). Under the Becker model of discrimination, there are several different types of discrimination, but the relevant part to the topic of this paper pertains to the variant directed at employees (Becker 1995). Retrospectively, it's important to note that domestic businesses were legally permitted to discriminate in both their hiring and payment processes until the passage of the Civil Rights Act of 1964 (Akee & Yuksel 2010). In South American countries like Ecuador, socioeconomic disparities based in racial discrimination have persisted since colonial times (Gallardo & Ñopo 2019). Workers who identify as black/African American or native American will most likely earn less than their white counterparts, and while Asian workers typically earn more than white workers, they are not necessarily exempt from discrimination either (Chang & Reed 2003). Researchers believe that 33 to 50% of African American men earn 12-15% less than their white counterparts (Akee & Yuksel 2010).Wage gaps also become increasingly severe at lower income levels, which makes it vital that the economic community closely examine industries like agriculture (Gallardo & Ñopo 2019). This effect also extends to the geographic spectrum, as ethnic minorities are likely to be concentrated in a specific area of a given city or state. This ties into the concept of spatial mismatch, as minority workers are more likely to live further away from high-paying job opportunities (Chang and Reed 2003). Also, in the event that wage gaps or wage discrimination are discovered in an employment sector, governmental regulation or union formation can mitigate its effects to a degree. Regarding the locations of the case studies serving as a basis for this research paper, South American countries make for ideal locations for studies of this nature, due to their high concentrations of populations of mixed African descent (Iowell 1995). On the domestic side of things, one of the case studies was conducted in California. The fact that evidence of race-based wage gaps was found in such a diverse state is an alarming omen for the rest of the United States (Chang & Reed 2003).

However, despite the significant evidence pointing to racial/ethnic discrimination, there are a variety of other factors that are necessary to take into account when looking at earnings. For one, education plays a significant role in determining one’s wage, and since ethnic minorities are likely to have lower education levels and less job experience, this can create a cycle of perceived discrimination (Gallardo & Ñopo 2019). Gaps in AFQT (Armed Forces Quantile Test; an exam that is utilized to gauge entrance eligibility into the Armed Services) scores can explain a decent portion of the white-Hispanic/Latino and white-black wage gaps of years past (Akee & Yuksel

2010). Ideally, under capitalist theory, once minorities are able to obtain equal educational opportunities, wage discrimination will disappear (Chang & Reed 2003). However, while this would stabilize Asian and Hispanic/Latino earnings, black workers would still be earning less than their white counterparts even after accounting for their schooling. Economic instability can also play a role, as higher unemployment rates will likely lead to greater wage disparities between various groups.

The gender divide can significantly impact income in a similar manner, so using interaction terms with a female indicator is highly recommended (Akee & Yuksel 2010). Gender wage gaps are highly comparable to ethnic ones, so this will be an important statistic to keep an eye on going forward. Age, marital status, immigrant status and number of children can all also influence one's wages. In terms of actually evaluating the information in the aforementioned case studies, linear regressions and Oaxaca-Blinder decompositions seem to be the two most common methods of examining wage gaps (Iowell 1995). In many of the papers, four primary racial/ethnic classifications were utilized: white, black/African American, Asian, and Hispanic/Latino (Chang & Reed 2003). Panel data following set individuals over long periods of their lives were frequently used. Despite their focus on discrimination against individuals of a singular race, there are some passing references to the consequences of profiling toward multiethnic groups. The wages of biracial (white and African American) workers were between those of solely white and solely black workers, but closer to solely black wages (Akee & Yuksel 2010). Part of the reason that multiethnic wage gaps are so difficult to track is that identity can be highly subjective and can even change over time. In Brazil, even a hint of African blood is enough to label an individual as completely black (Iowell 1995). Racial classification is also heavily dependent on one's social class, as upper-class individuals of mixed African descent are more likely to identify as mulatto or white. However, despite all of these potential complications, it has been shown that those of mixed Afro-Hispanic descent earned significantly less than their white counterparts (Iowell 1995).

Dataset

Using all of these previous studies on race-based wage gaps as a basis for the model, next a dataset was needed that allows for multiracial/multiethnic self-identification by its respondents. The National Longitudinal Survey of Youth 1979 fulfills this prerequisite and possesses a multitude of essential control variables as well. In the NLSY, respondents are initially classified as Hispanic/Latino or non-Hispanic/Latino (per the Federal guidelines for surveys of this nature), and are then split into different racial categories. The panel data follows 12,686 people born between 1957 and 1964. The data collection from the year 2002 allowed candidates to select multiple ethnic identities. Part of the reason that multiracial identities are so hard to track is that, in this case, only 24% of identifications were consistent over the course of the survey, so only one round of responses will be used for the sake of consistency (Light & Nandi 2007). Looking at the spread of the ethnic identification groups (which will be numerically detailed in the Methodology section), the white multiracial combinations were far more common than the non-white combinations. White and Hispanic data are also heavily correlated (likely due to many individuals denoting white as their race and Hispanic/Latino as their ethnicity), something that will most likely carry over into their racial combination (Light & Nandi 2007). All of the possible racial/ethnic identity options were selected as explanatory variables in a customized dataset created for use during the experiment, along with controls for sex, age, education, AFQT test scores, marital status, and number of children. These control variables will serve as proxies for worker productivity. Below is a table containing the averages and standard deviations of the quantifiable non-race variables.

Summary Statistics

	(1)
Children	0.965
	(1.236)
Income	13036.4
	(27002.9)
Education	12.56
	(2.421)
Female	0.495
	(0.500)
Age	33.34
	(11.23)
Potential Experience	14.42
	(11.59)
Potential Experience Squared	342.3
	(407.0)
Logged Earned Income	9.352
	(1.405)
AFQT Composite Score	2510.1
	(1068.0)
Start Date	965.3
	(44.32)
Married	0.442
	(0.497)
Observations	342522

mean coefficients; sd in parentheses

However, before said tests could be implemented, there were a vast variety of modifications that needed to be made to the data in order to move forward. First, a dummy variable representing the female portion of the sample was derived from the pre-established *Sex* variable. Next, observations of each participant's ages were generated for each of the interview years using the 1979 Age variable as a base. While the NLSY dataset did include variable that observed each survey participant's age in the necessary years, 1979 was the only year where all of the participants responded, which mean that using all of the age variables would result in right-skewed results as less people responded over time. The Highest Grade Completed/Education variable suffered a similar problem of progressively lower response rates as the data approaches the present day, as participants would stop responding once they finished their educations. To adjust for this, the variables' most recent response will be continuously utilized in regressions, serving as a constant reference for those who have completed their education. Potential work experience was also implemented into the runs of the regression a modified estimate derived from a Start Date miniature dataset[1], which was created from the NLSY79 response dates. This dataset tracks how many months have elapsed since the survey respondents have exited school, provided that said exit lasts for at least a year. The potential experience variable also has a squared value that is regressed as a separate variable. Next, a representative variable of one's Armed Forces Qualification Test (AFQT) score is constructed using the official military method of evaluation on the sub-scores provided in the NLSY dataset. Additionally, the Number of Children and Married, which is a simplified dummy variable derived from the Marital Status variable, were both included in the regressions, along with interaction terms for both of the aforementioned variables with the Female variable in order to account for gender disparities in those areas. Finally, a variable derived from the dataset that

1 The Start Date dataset was created and supplied by Dr. Audrey Light of the Department of Economics at The Ohio State University

accounts for annual income will serve as the response variable of the regressions and will be transformed using the "log" command in order to evaluate the earnings differences between the different racial/ethnic groups on a percentage-based scale. Controls for each year will also be implemented in order to control for inflation. The standard errors of each regression will be clustered by the ID Codes of each survey respondent.

Methodology

Once the customized dataset was created, it was imported into Stata, a statistical data analysis software. Both linear regression models and F-statistic tests were conducted in order to determine whether the variables representing the different multiethnic combinations are statistically significant or not. In order to create variables that effectively represent all possible combinations of multiethnic identities, some groups in the data had to be reorganized into four primary categories of White, Black/African American, Asian and Other. Those who identified as Hawaiian/Pacific Islander and American Indian were categorized as part of the Asian and Other groups respectively. Once the four primary race variables were constructed, all possible combinations were multiplied together to create dummy interaction variables for each multiethnic group. Immediately after this, the interaction terms were subtracted from the singular race variables in order to create variables representative of those who solely identify as one race/ethnicity. A control variable for anyone who identifies as part of the Hispanic/Latino ethnicity will also be included. Below is a table containing the frequencies at which respondents identified with different racial and ethnic classifications.

Race/Ethnicity	Frequency
White	4,531
White Only	4,461
Black	2,333
Black Only	2,289
Hispanic	1,412
Hispanic Only	471
Asian	44
Asian Only	33
Other	445
Other Only	364
Modified Other Only	111
White & Black	18
White & Hispanic	631
White & Asian	10
White & Other	59
White & Modified Other	65
Black & Hispanic	26
Black & Asian	8
Black & Other	35
Black & Modified Other	39
Asian & Other	4
Hispanic & Modified Other	306

The first regression run will serve as a reference; containing only the basic race variables (not the "Only" variations, which allows for overlap/multiple selections from an individual) and control variables. Once the regression results have been computed by STATA, an F test will be conducted on the three racial terms to evaluate whether their inclusion is necessary to the regression. Below is the model equation:

Log(Earned Income) = β_0 + β_1(Black) + β_2(Asian) + β_3(Other) + β_4(Hispanic/Latino Ethnicity) + β_5(Female) + β_6(Education) + β_7(Potential Experience) + β_8(Potential Experience Squared) + β_9(AFQT Score) + β_{10}(Married) + β_{11}(Children) + β_{12}(Female x Married) + β_{13}(Female x Children) + β_i(i.Year) + ε

The second run substitutes the basic race variables with the "Only" variations and adds the racial interaction terms in order to account for multiracial identification. The model equation is as follows:

Log(Earned Income) = β_0 + β_1(Black Only) + β_2(Asian Only) + β_3(Other Only) + β_4(White & Black) + β_5(White & Asian) + β_6(White & Other) + β_7(Black & Asian) + β_8(Black & Other) + β_9(Asian & Other) + β_{10}(Hispanic/Latino Ethnicity) + β_{11}(Female) + β_{12}(Education) + β_{13}(Potential Experience) + β_{14}(Potential Experience Squared) + β_{15}(AFQT Score) + β_{16}(Married) + β_{17}(Children) + β_{18}(Female x Married) + β_{19}(Female x Children) + β_i(i.Year) + ε

For the third run of the regression, the Asian and Other categories were combined due to low response counts from the two groups, resulting in the following model:

Log(Earned Income) = β_0 + β_1(Black Only) + β_2(Other Only) + β_3(White & Black) + β_4(White & Other) + β_5(Black & Other) + β_6(Hispanic/Latino Ethnicity) + β_7(Female) + β_8(Education) + β_9(Potential Experience) + β_{10}(Potential Experience Squared) + β_{11}(AFQT Score) + β_{12}(Married) + β_{13}(Children) + β_{14}(Female x Married) + β_{15}(Female x Children) + β_i(i.Year) + ε

For the fourth and final run of the regression, Hispanic/Latino is used as a primary racial/intersectional group. While the Hispanic/Latino group isn't typically considered a race, this experiment pertains more to the concept of self-identity than pre-established racial groups. As a result of this group's prominence within the ethnic makeup of the United States, it is the sole ethnicity to be included in the regressions with its own designated intersectional categories. After creating new combinations, the following regression model is used:

Log(Earned Income) = β_0 + β_1(Black Only) + β_2(Hispanic Only) + β_3(Other Only) + β_4(White & Black) + β_5(White & Hispanic) + β_6(White & Other) + β_7(Black & Hispanic) + β_8(Black & Other) + β_9(Hispanic & Other) + β_{11}(Female) + β_{12}(Education) + β_{13}(Potential Experience) + β_{14}(Potential Experience Squared) + β_{15}(AFQT Score) + β_{16}(Married) + β_{17}(Children) + β_{18}(Female x Married) + β_{19}(Female x Children) + β_i(i.Year) + ε

After running all four regressions, the significance of the multiracial/multiethnic variables were evaluated using t-tests and F-tests. Said F-tests were conducted on the multiracial interaction terms in all runs except the first. The coefficients were also analyzed in order to determine which of the single race categories the multiethnic wages tended to sway toward.

Results

Below are the results of the four regressions detailed above. It’s important to note that a white male who is unmarried and has no children serves as the reference relative to the coefficients of the explanatory variables detailed below.

	(1)	(2)	(3)	(4)
	Logged Earned Income	Logged Earned Income	Logged Earned Income	Logged Earned Income
Black (Raw Data)	-0.185***			
	(0.0152)			
Asian (Raw Data)	-0.00923			
	(0.108)			
Other (Raw Data)	-0.0718**			
	(0.0321)			
Black Only		-0.187***	-0.186***	-0.186***
		(0.0153)	(0.0154)	(0.0153)
Hispanic Only (Intersectional)				0.0417
				(0.0266)
Asian Only		-0.0979		
		(0.135)		
Other Only		-0.0511	-0.135*	-0.135*
		(0.0362)	(0.0701)	(0.0701)
White & Black		0.248	0.283**	0.299**
		(0.156)	(0.142)	(0.143)
White & Asian		0.237		
		(0.212)		
White & Other		-0.153**	-0.0975	-0.112
		(0.0727)	(0.0705)	(0.0710)
White & Hispanic				0.0640**
				(0.0240)

Black & Asian		0.0745		
		(0.160)		
Black & Other		-0.337**	-0.257**	-0.252**
		(0.136)	(0.116)	(0.116)
Black & Hispanic				-0.233
				(0.163)
Asian & Other		0.151		
		(0.354)		
Hispanic & Other				0.0387
				(0.0359)
Hispanic (Ethnicity)	0.0633***	0.0603***	0.0362**	
	(0.0180)	(0.0182)	(0.0174)	
Female	-0.140***	-0.141***	-0.141***	-0.140***
	(0.0154)	(0.0154)	(0.0154)	(0.0154)
Education	0.128***	0.129***	0.129***	0.129***
	(0.00312)	(0.00312)	(0.00312)	(0.00312)
Potential Experience	0.0695***	0.0695***	0.0696***	0.0696***
	(0.00283)	(0.00282)	(0.00282)	(0.00282)
Potential Experience Squared	-0.00131***	-0.00131***	-0.00131***	-0.00131***
	(0.0000666)	(0.0000665)	(0.0000665)	(0.0000665)
AFQT Composite Score	0.0000730***	0.0000732***	0.0000735***	0.0000878***
	(0.00000694)	(0.00000696)	(0.00000695)	(0.00000696)
Married	0.422***	0.422***	0.422***	0.422***
	(0.0134)	(0.0133)	(0.0134)	(0.0133)

Children	0.0142**	0.0143**	0.0144**	0.0142**
	(0.00655)	(0.00654)	(0.00655)	(0.00655)
Female * Married	-0.395***	-0.395***	-0.394***	-0.395***
	(0.0191)	(0.0191)	(0.0191)	(0.0191)
Female * Children	-0.151***	-0.151***	-0.152***	-0.152***
	(0.00881)	(0.00881)	(0.00881)	(0.00880)
Constant	5.947***	5.945***	5.945***	5.945***
	(0.0409)	(0.0410)	(0.0411)	(0.0411)
N	189105	189105	189105	189105
R^2	0.480	0.480	0.480	0.480
F-Statistic	51.26	2.56	2.71	3.12
F-Probability	0***	0.0178**	0.0436**	0.0046***

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Based on the regression and F-test results displayed above, it appears that there have been some forms of wage gaps for individuals of more than one race at some point in the past four deacdes. As established by previous literature on this subject, all of the regressions point to staistically significant wage gaps for those who solely identify as black or African American. In the alternative models where the Other and Asian categories are combined, there were also significant wage gaps for individuals who solely identified as being part of this category. Taking this background information into account, it comes as no surprise that the two multiethnic groups to exhibit significant numerical evidence of wage gaps are “White & Other” and “Black and Other”. All three of the unique[2] regressions to this paper returned highly signficant F-tests when compared with models that lack all of the mutliethnic variables; proving that their inclusion is

[2]This refers to the three rightmost columns of the regression table that contain the multiracial terms, or Columns 2,3 and 4

beneficial toward understanding the data. However, it is imperative to note that the cause of these wage gaps cannot be explicitly determined by the findings of this paper.

Regarding which of the two ethnicities the multiracial wages tend to lean toward, the outcomes were mostly as initially predicted by the background literature. All of the Asian combinations tended to lean closer to white wages than those of solely Asian wages. For the most part, the earnings of those in mixed Hispanic/Latino and "Other" (meaning those of either native American or American Indian descent) categories were closer to those of their respective races than those of solely white workers. Finally, and surprisingly, the Black/White and Black/Asian combinations had wages that were closer to sole white earnings than their respetive single minority earnings. However, this can be attributed to low sample sizes of the combinations, coming in at 18 and 8 respondents respectively. The results for Black/Hispanic and Black/Other pivoted toward solely Black earnings as expected.

Possible Expansions

While the results of the regressions were more than satisfactory, there are a myriad of potential tecchniques that could expand on this concept, given the proper conditions and a substantial sample size. For example, adding control variables that divide the job market into different sectors/industries could allow for an in-depth examination of the differing wage disparities for the respective ethnic groups (Akee & Yuksel 2010). Creating interaction terms for each multiethnic combination crossed with each year could generate a similar effect. While the focus of this experiment was using regressions in order to quantify wage disparities between different ethnic groups, using the Oaxaca-Blinder decomposition method in order to determine exactly how much of the aforementioned wage gaps are due to observed differences in human capital versus other unobserved factors, which could potentially include discrimination. This could serve as an indicator of progress regarding the elimination of discrimination from the American workplace (Gallardo & Ñopo 2019). Adding a control grouping mechanism for immigrant status could help to determine if immigrants of minority ethnicities are being domestically discriminated against (Chang & Reed v). Finally, using hourly and weekly income as the response variable of the linear regressions could supplement the annual earnings regression described in this experiment well, although including assets earned under those models could become difficult.

Conclusion

In conclusion, there is significant statistical evidence that suggests that a wage gap exists for individuals who identify as either Caucasian and Native American (or other indigenous groups) or African American and Native American/Asian. While these variables don't possess large amounts of explanatory power when it comes to factors like the regression correlation coefficient, they are highly significant variables and merit further study. The experiment procedure also validated the initial hypothesis that being of Hispanic/Latino, African American, or Native American descent indicates that these workers' incomes will be pulled in the direction of the "uniform" wages of these three races/ethnicities. The F-tests conducted on all of the unique regressions created during this experiment returned incredibly high values, signifying that the inclusion of multiethnic terms increases the accuracy of their results. Now that the existence of wage gaps for multiracial populations in the domestic workplace has been confirmed, the next step would be to determine the cause of this phenomenon using various economic and statistical techniques. While not much can be done to combat discrimination outside of creating and updating legislation, domestic educational and economic welfare programs may need to be reexamined in the event that differences in education or skill are to blame for the wage gaps occurring. While society may never truly be rid of inequality and biases, the more the world can determine and begin to understand the sources of disparities and discrimination, the more these problems can be minimized in the workplace and beyond.

Bibliography

"Access Data / Investigator." *Access Data / Investigator | National Longitudinal Surveys*, 2019,www.nlsinfo.org/content/access-data-investigator.

Akee, Randall, and Mutlu Yuksel. "Skin Tone's Decreasing Importance on Employment: Evidence from a Longitudinal Dataset, 1985-2000." *Econstor*, 2010, www.econstor.eu/bitstream/10419/44175/1/64452880X.pdf.

Becker, Gary Stanley. *The Economics of Discrimination*. Univ. of Chicago Press, 1995.

Benjamin, Tia. "The Ways Discrimination Negatively Affects Businesses." *Small Business – Chron.com*, Chron.com, 26 Oct. 2016, smallbusiness.chron.com/ways-discrimination-negatively-affects-businesses-36925.html.

Chang, Jennifer and Deborah Reed. "Racial and Ethnic Wage Gaps in the California Labor Market." 2003. media.asian-nation.org/racial-ethnic-wage-gaps-ca.pdf.

Gallardo, Lourdes, and Hugo Ñopo. "Ethnic and Gender Wage Gaps in Ecuador." *Econstor* , 19 Sept. 2019, www.econstor.eu/bitstream/10419/51500/1/601743776.pdf.

Iowell, Peggy. "Race, Gender, and Development in Brazil." *Lua Nova: Revista De Cultura e Política*, CEDEC, 1995, www.scielo.br/scielo.php?pid=S0102-64451995000100003&script=sci_arttext&tlng=es.

Light, Audrey, and Alita Nandi. "Identifying Race and Ethnicity in the 1979 National

Longitudinal Survey of Youth.” *SpringerLink*, Kluwer Academic Publishers, 6 Mar. 2007, link.springer.com/article/10.1007/s11113-007-9021-1